Journal of Technology Research (JTR)
Volume 3, Special Issue, (2025), pp 515-525 ISSN 3005-639X
The 7th Conference of Engineering Sciences and Technology
(CEST-2025)

Hybrid Stacking Ensemble Model for Phishing URL Detection Using PCA and Machine Learning

Mohammed M. Elsheh[1], Ebtisam Abolawaifa[1]*

[1]Computer Science, The Libyan Academy for Postgraduate Studies, Misurata, Libya.

*Corresponding author email: m.elsheh@lam.edu.ly

Received: 22-09-2025 | Accepted: 25-11-2025 | Available online: 25-12-2025 | DOI:10.26629/jtr.2025.48

ABSTRACT

The rapid growth of internet usage has transformed cybercrime into a formidable global challenge. Among digital threats, phishing stands out as particularly dangerous due to its deceptive approach. Cybercriminals employ fake yet convincing URLs to steal users' sensitive information, causing significant financial and personal damage. This escalating threat demands advanced countermeasures. This study addresses this critical need by proposing a hybrid machine learning (ML) model specifically designed to improve malicious URL identification. The innovative model integrates three powerful algorithms: Logistic Regression (LR), Artificial Neural Networks (ANN), and Random Forest (RF). These are combined within an advanced stacking ensemble architecture that strategically leverages each algorithm's unique analytical capabilities. This multi-layered approach enables comprehensive threat analysis from different perspectives. To optimize model efficiency and performance, we implemented Principal Component Analysis (PCA) for intelligent feature selection, ensuring optimal computational resource utilization. Our research utilized a substantial dataset of over 11,000 carefully labelled URLs sourced from Kaggle. The dataset underwent meticulous preparation, including appropriate balancing techniques to mitigate class imbalance issues that could compromise model accuracy. Through rigorous evaluation using key performance metrics accuracy, precision, recall, and F1-score the model demonstrated exceptional efficacy. Remarkably, the hybrid ensemble achieved an outstanding accuracy of approximately 99.55%, significantly surpassing all individual base models. This superior performance highlights the model's strong potential for immediate deployment in real-time phishing detection systems. It offers organizations a proactive and reliable defence mechanism in the ongoing battle against evolving cyber threats, representing a significant advancement in cybersecurity protection for today's sophisticated digital landscape.

Keywords: Phishing URLs, Voting classifier, Stacking classifier, Logistic Regression, Artificial Neural Network, Random Forest, Cyber Security.

نموذج التجميع الهجين للكشف عن عناوين URL الاحتيالية باستخدام تحليل المكونات الرئيسية (PCA) والتعلم الآلي

محمد مصباح الشح[1]، ابتسام إبراهيم أبولويفة[1]

[1]علوم الحاسوب، الاكاديمية الليبية للدراسات العليا، مصراتة، ليبيا.

ملخص البحث

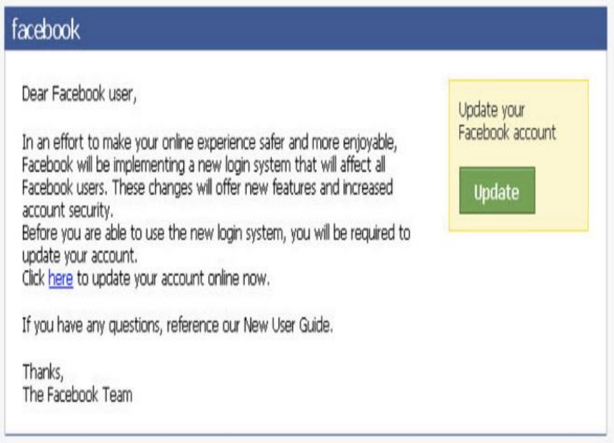
شكل الانتشار السريع للإنترنت تحديًا عالميًا خطيرًا في مواجهة الجرائم الإلكترونية. وتعتبر هجمات التصيد الاحتيالي من أخطر هذه التهديدات بسبب طبيعتها الخادعة، حيث يستخدم المجرمون روابط مزيفة لكنها تبدو حقيقية لسرقة المعلومات الحساسة للمستخدمين. يتطلب هذا التهديد المتصاعد تدابير مضادة متطورة. تستجيب هذه الدراسة لهذا التحدي من خلال تطوير نموذج هجين يجمع بين ثلاث تقنيات ذكية هي Logistic Regression (LR) و Artificial Neural Networks (ANN) و Random Forest (RF) ضمن هيكلية تجميع متطورة. تُدمج هذه الخوارزميات ضمن بنية تكديس مُتقدمة تُوظّف بشكل استراتيجي القدرات التحليلية الفريدة لكل خوارزمية. يُتيح هذا النهج متعدد الطبقات تحليلًا شاملًا للتهديدات من وجهات نظر مُختلفة. لضمان كفاءة النظام، تم استخدام تقنية تحليل المكونات الرئيسية (PCA) لاختيار أهم السمات، كما تم تدريبه على مجموعة شاملة تحتوي على أكثر من 11,000 موقع إلكتروني مصنف. وتم تطبيق تقنيات موازنة متقدمة لمعالجة عدم التوازن في البيانات. أظهر التقييم الشامل باستخدام مقاييس accuracy و precision و recall و F1-score أداءً استثنائيًا للنموذج، حيث حقق دقة 99.55% متفوقًا على النماذج الفردية. يُبرز هذا الأداء المتفوق الإمكانات القوية للنموذج للنشر الفوري في أنظمة الكشف عن التصيد الاحتيالي في الوقت الفعلي. كما يوفر للمؤسسات آلية دفاع استباقية وموثوقة في المعركة المستمرة ضد التهديدات السيبرانية المتطورة، مما يمثل تقدمًا كبيرًا في حماية الأمن السيبراني في ظل البيئة الرقمية المتطورة اليوم.

الكلمات الدالة: عناوين URL الاحتيالية، مصنف التصويت، مصنف التكديس، الانحدار اللوجستي، الشبكة العصبية الاصطناعية، الغابة العشوائية، الأمن السيبراني

INTRODUCTION

In the contemporary digital world, phishing attacks have emerged as one of the most pervasive and detrimental forms of cybercrime. These attacks predominantly exploit malicious URLs. These URLs redirect unsuspecting users to counterfeit websites. These websites are meticulously crafted to resemble legitimate platforms. The primary objective of such attacks is to illegally obtain sensitive users' information. This includes authentication credentials, financial details, and other personal data. Cybercriminals increasingly employ advanced evasion techniques. One common technique is the use of deceptive subdomains embedding well-known brand names. Another way is the use of URL shortening services to conceal destination addresses. They also deploy HTTPS protocols to create a false sense of trust. Additionally, attackers often replicate the visual design of authentic websites. They mimic user interfaces as well. These strategies further enhance the deceptive potential of phishing attacks.

For instance, Figure 1 shows that phishing can be perpetrated by tricking someone into clicking a malicious link that seems legitimate.



Dear Facebook user,

In an effort to make your online experience safer and more enjoyable, Facebook will be implementing a new login system that will affect all Facebook users. These changes will offer new features and increased account security.
Before you are able to use the new login system, you will be required to update your account.
Click here to update your account online now.

If you have any questions, reference our New User Guide.

Thanks,
The Facebook Team

Fig 1. An example of a phishing attack [1]

This approach is used instead of trying to break through a computer's defence systems. The malicious links are placed within the body of the message. They are designed to make it appear that they go to the spoofed organization. This is done by using that organization's logos and other legitimate content [1].

However, phishing reached record level in 2023, marking the worst year on record for this type of attack. Phishing incidents saw a sharp rise in the second half of the year. This came after a slight dip in the second quarter. The Anti-Phishing Working Group (APWG) recorded just over 1 million unique phishing attacks in Q4 alone. Overall, they witnessed almost 5 million such incidents throughout the full year [2].

Given the dynamic and evolving nature of phishing techniques, traditional rule based detection systems have proven insufficient [3]. In response, machine learning (ML) has emerged as a promising approach, capable of analyzing a wide array of URL based features such as URL length, the presence of abnormal characters, domain registration attributes, and phishing-related lexical patterns [4]. By identifying latent patterns and correlations within these features, ML models can effectively distinguish between benign and malicious URLs [5]. Unlike static detection methods, ML based systems possess the adaptive capacity to recognize novel and previously unseen attack vectors, thereby enhancing detection accuracy and reducing false positive rates [6]. Feature selection in classification models is typically done using filter, wrapper, or embedded methods [7]. In this study, PCA was used as a filter method to reduce dimensionality and keep the most relevant features [8][9] . Achieving high classification accuracy mainly relied on tuning key hyperparameters like increasing the number of trees in the RF [10], modifying the hidden layers and learning rate in the ANN [11] and adjusting the regularization strength and solver in LR which collectively improved the model's performance [12].

1. RELATED WORK

Phishing detection has gained increasing attention in recent years due to the growing sophistication of cyberattacks and the limitations of traditional security mechanisms. Various ML and hybrid approaches have been proposed in the literature to address this challenge. Below is a summary of notable studies in this field.

In 2021, a group of researchers developed a classification framework combining ANNs and RF with ensemble techniques. They used a Kaggle dataset of 11,055 URLs. The models achieved strong results, with ANN reaching 98.72% accuracy and 100% precision and recall for the phishing class, while RF recorded 97.65% of accuracy [13].

In 2022, a study proposed a hybrid feature-based detection model by extracting 15 URL-related and 10 hyperlink-related features, avoiding dependence on third-party services. It was evaluated on a balanced dataset of 6,000 URLs. The XGBoost classifier gained the best result, achieving 99.17% accuracy, 98.81% recall, and a low false positive rate [14].

In a 2023, another study suggested, a hybrid model that integrates SVM with Ant Colony Optimization (ACO) and Deep Belief Networks (DBN). This approach was evaluated on 12,000 labelled samples, resulting in 97.54% accuracy and improved performance across all evaluation metrics compared to standard SVM classifiers [15].

In the same year, another study proposed a hybrid ensemble model (LSD) that combines LR, SVM, and DT. The model is trained on 11,054 URLs from Kaggle. The LSD model achieved an accuracy of 98.12%, outperforming individual classifiers [16].

Also in 2023, a deep learning model combining Deep Neural Networks (DNN) with Bidirectional Long Short-Term Memory (BiLSTM) was developed to **enhance phishing detection by leveraging both semantic information and sequential patterns in URLs.** The architecture model fused NLP features and character level embeddings, achieving accuracy of 99.21% and 98.79% on

PhishTank and Ebbu2017 datasets, respectively[17].

A comparative evaluation in 2023 explored several ML classifiers, including DTs, RF, AdaBoost, KNN, SGD, Extra Trees, and NB, using the ISCX-URL-2016 dataset with over 650,000 URLs. Extra Trees and RF achieved the best results, both exceeding 91% accuracy [6].

Recently in 2024, a hybrid feature selection approach combining Mutual Information Gain with Genetic Algorithms (GN) was introduced to improve phishing detection in IoT environments. The method was evaluated on a 10,000 record dataset using XGBoost, achieving an accuracy of 98.3% and a precision of 99% [18].

Another study in 2024 focused on enhancing phishing detection using a GN based feature selection method. On a dataset containing over 87,000 URLs from PhishTank, the RF achieved 92.93% accuracy and 89.05% recall using 44 optimized features [19].

Also in 2024, a novel model called ResMLP was proposed, combining residual pipelining with multi-layer perceptron networks. It was trained on over 500,000 URLs from Kaggle. The model achieved 98.29% accuracy, 98.10% precision, and 98.94% F1-score, showing promise for real-time phishing detection [20].

In 2025, a hybrid ML approach was developed for phishing website detection using URL-based features. The model was trained on the UCI dataset containing 11,055 instances and evaluated with multiple classifiers, including DT, RF, SVM, and AdaBoost. RF achieved the best performance with 97.7% accuracy, 99% precision, and 97% F1-score, consistently outperforming the other algorithms [21].

Also in 2025, a study proposed a hybrid stacking model called PhishDef-XRL, which combined RF and XGBoost with LR as a meta-classifier. When evaluated on 88,647 URLs, the model achieved 97.16% accuracy, outperforming both base models and other hybrid approaches in phishing URL detection [22].

Recently, in 2025, a study proposed the EGSO-CNN model, which integrated Variational Autoencoders for feature extraction and an Enhanced Grid Search for optimization, when evaluated on a custom-built dataset of 27,534 URLs, it achieved a high accuracy of 99.44% for phishing URL detection [23].

In the same year, a study proposed the PDSMV3-DCRNN ensemble model, which employed a Conditional Wasserstein GAN (CWGAN) to handle data imbalance and the Binary Grey Goose Optimization Algorithm (BGGOA) for feature selection. When it was evaluated on four benchmark datasets, including ISCX-URL-2016 and Mendeley_2020, the model achieved a high accuracy of 99.21% with a fast-training time of 0.11 seconds, outperforming existing methods in both speed and detection performance [24].

Finally, in 2025, another study suggested an approach to using several deep learning models, including Feedforward Neural Network (FNN), DNN, Wide and Deep, and TabNet. The models were trained on the Mendeley 2020 dataset 58,645 URLs with 111 features, with feature selection performed via Permutation Importance to reduce the feature set to just 14 key attributes. The FNN model emerged as the most efficient, achieving a high accuracy of 94.46% and the best anti-phishing score, demonstrating a robust balance between performance and computational cost for real-time detection [25].

As it can be noticed from all previous studies, the issue of phishing is still a problem that threatens all Internet users around the world, so the field is still open for more research to solve and confront this scourge.

2. RESEARCH DESIGN

The framework of this study consists of four main phases as illustrated in Figure 2:

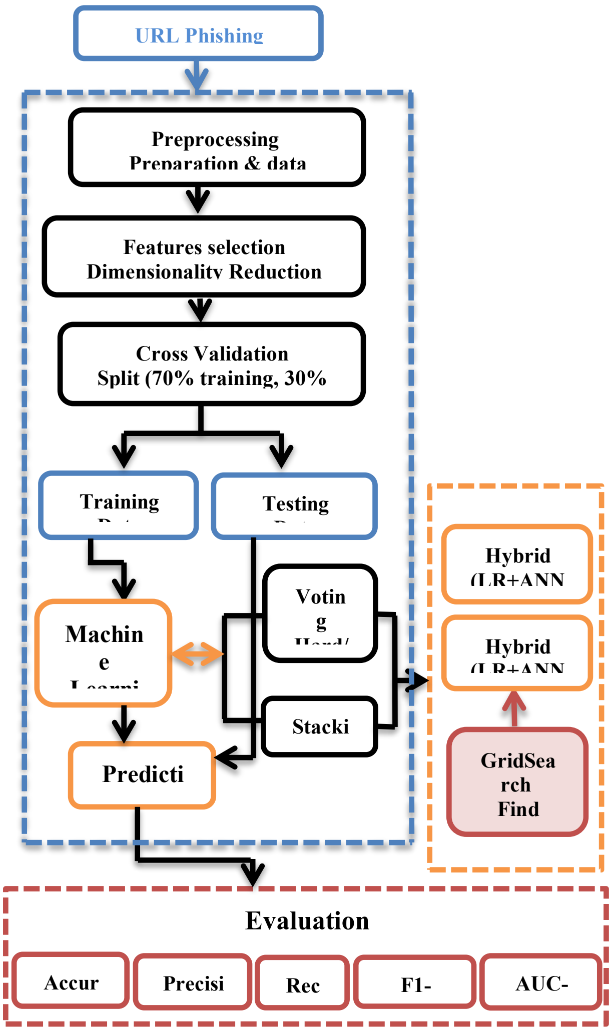


Fig 2. The framework of the Study.

3.1 Dataset Collection

The dataset was collected from the well-known dataset repository Kaggle and stored as a CSV file, which provides benchmark datasets for research purposes. It consisted of 11,054 records and 32 attributes extracted from over 11,000 URLs [26].

The dataset consisted of two classes: phishing URLs (label 1) and legitimate URLs (label 0), as illustrated in Figure 3. After removing null values and refining the feature vectors, the class imbalance between the majority (legitimate) and minority (phishing) classes was addressed by employing under sampling[27]. This technique ensured equitable representation of both classes and prevented the model from developing a bias towards the dominant class. The resulting balanced corpus was then split into a 70% training set and a 30% test set using cross-validation. The training set was used to build the ML model, while the test set was reserved for its evaluation.

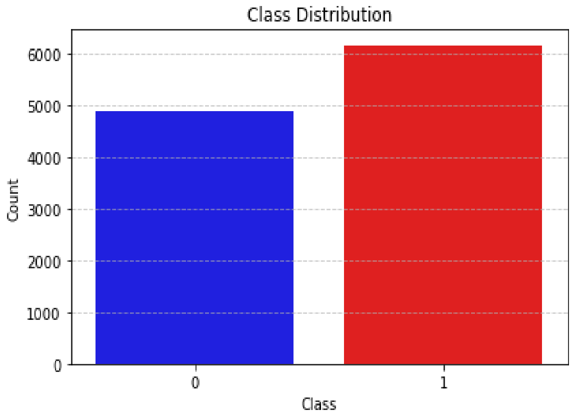


Fig 3. Dataset presentation according to the number of class labels.

3.1.1 Features Selection

In this phase, the feature selection process was performed using **PCA technique**. It was applied as a dimensionality reduction technique. It is an unsupervised statistical method that transforms high-dimensional data into a lower dimensional subspace[28]. This is done by identifying orthogonal principal components that capture the maximum variance in the data. The transformation relies on eigen decomposition of the covariance matrix. As a result, PCA produces new, uncorrelated features that are linear combinations of the original attributes. These components retain the most important patterns in the dataset. By removing low-variance components, it reduces both noise and computational complexity. However, the abstract nature of the PCA may reduce the interpretability of individual features. Despite this, PCA preserves the most informative aspects of the data needed for classification. Its use in this context improves learning by filtering redundant features and enhancing the model's generalization capability [9].

3.1.2 Classification

At this phase, a hybrid classification model combining (LR, ANN, and RF) was constructed using both Voting (hard/soft) and Stacking ensemble techniques. To enhance the predictive performance of each base classifier, hyperparameter tuning was performed using GridSearch Cross Validation (CV), which systematically explores the optimal values for the model parameters to perform the classification process[28].

3.1.3 Evaluation

The final phase is the evaluation of the proposed approach. The classification process was executed in four steps: Voting before tuning, voting after tuning, stacking before tuning, and stacking after tuning. The performance of each step was evaluated using standard classification metrics, including Accuracy, Precision, Recall, F1-score, and AUC-ROC, to ensure a comprehensive and reliable comparison of model effectiveness before and after optimization [29].

3. RESULTS AND DISCUSSION

As a result of applying PCA, fifteen features were identified as the most significant based on their low reconstruction errors.

4.1 Hyperparameter Tuning

The hyperparameter tuning process for the base classifiers in the ensemble model was designed to optimize performance by balancing bias and variance[30]. LR, the regularization strength C was varied from 0.001 to 100, and the solvers `lbfgs` and `liblinear` were chosen for their suitability in binary classification tasks. In the ANN, different hidden layer configurations [(50), (100), (50,50)] and initial learning rates [0.001 and 0.01] were tested to assess learning depth and convergence behavior. For RF, the number of trees [100, 200, 500] and maximum depth [10, 20, None] were explored to manage complexity and reduce overfitting. These values were selected systematically to maximize cross-validation accuracy.

4.2 Hybrid Model (LR+ANN+RF) Using Voting Method

The ensemble model was evaluated under two setups: using default parameters and after fine-tuning with GridSearchCV. With default settings, soft voting achieved a cross-validation accuracy of 99.15%, precision of 99.24%, recall of 99.31%, and an F1-score of 99.27%. On another hand, hard voting slightly outperformed it, reaching 99.51% accuracy and an F1-score of 99.45%. After fine-tuning, soft voting yielded 99.14% accuracy, 99.44% precision, 99.24% recall, and a 99.34% F1-score. However, hard voting maintained strong and balanced results, with all metrics at 99.45%. These findings demonstrate the ensemble model's consistently and highly effective performance in phishing URL detection, regardless of tuning or voting strategy.

The confusion matrices for the (LR + ANN + RF) ensemble model, using both soft and hard voting before and after parameter optimization, revealed highly accurate classification performance with minimal misclassifications. When using default configuration, soft voting correctly identified 1,484 legitimate (TN) and 1,433 phishing (TP) instances, with 11 false positives (FP) and 11 false negatives (FN), while hard voting slightly improved results with 1,487 TN, 1,436 TP, and only 8 FP and 8 FN were detected as shown in Figure 4 and 5. After tuning parameters, soft voting maintained strong results with 1,485 TN, 1,434 TP, and 10 FP/FN each, and hard voting showed the best performance with 1,488 TN, 1,436 TP, 7 FP, and 8 FN, as shown in Figure 6 and 7. These results confirm the model's reliability and robustness in minimizing classification errors across all evaluation scenarios.

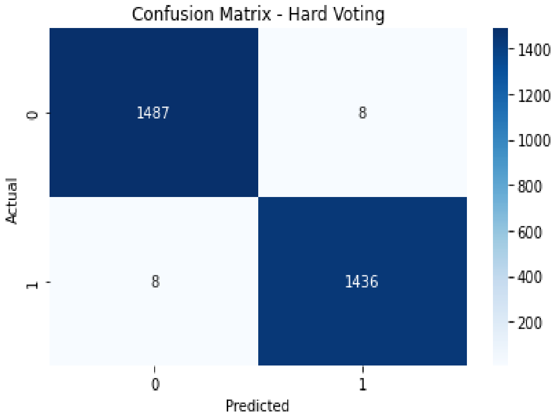


Fig 4. Confusion matrix for the model using hard voting with default parameters.

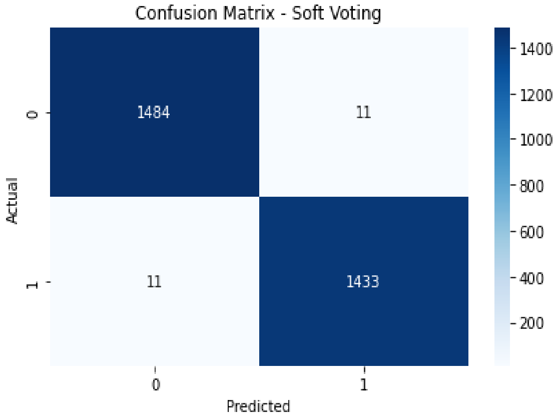


Fig 5. Confusion matrix for the model using soft voting with default parameters.

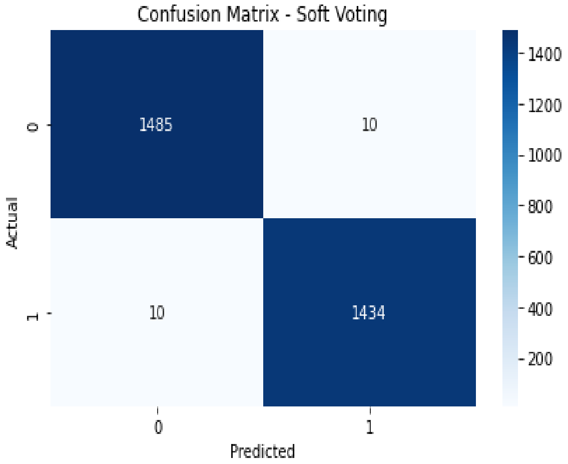


Fig 6. Confusion matrix for the model using soft voting after parameter tuning.

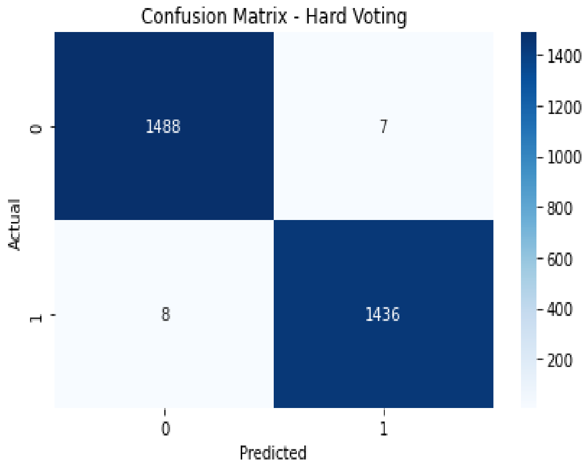


Fig 7. Confusion matrix for the model using hard voting after parameter tuning.

4.3 Hybrid Model (LR+ANN+RF) Using Stacking Method

The stacking ensemble model demonstrated exceptional performance in phishing detection both before and after hyperparameter tuning. With default parameters, it achieved a cross-validation accuracy of 99.52% and precision, recall, and F1-score of 99.65%, along with an AUC-ROC of 99.66%. After fine-tuning, the model slightly improved with a cross-validation accuracy of 99.55%, maintaining the same high precision, recall, F1-score 99.65%, and AUC-ROC 99.66%. These consistent results confirm the robustness and strong generalization ability of the stacking approach, even in the absence of parameter optimization.

The confusion matrices for the stacking model, before and after parameter tuning, show high classification accuracy. Without tuning as shown in Figure 8, the model correctly identified 1,491 legitimate (TN) and 1,437 phishing (TP) samples, with only 4 false positives (FP) and 7 false negatives (FN). After tuning as presented in Figure 9, it maintained 1,491 TN and improved to 1,439 TP, reducing false negatives to 5, while false positives remained at 4. These results reflect consistent and accurate performance across both configurations.

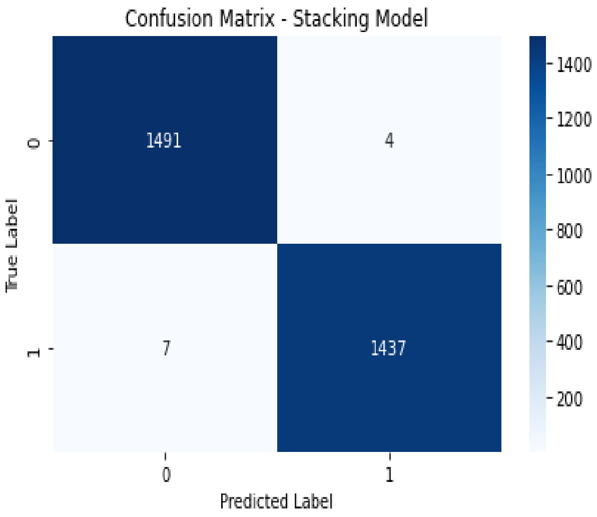


Fig 8. Confusion matrix for the model using stacking with default parameters.

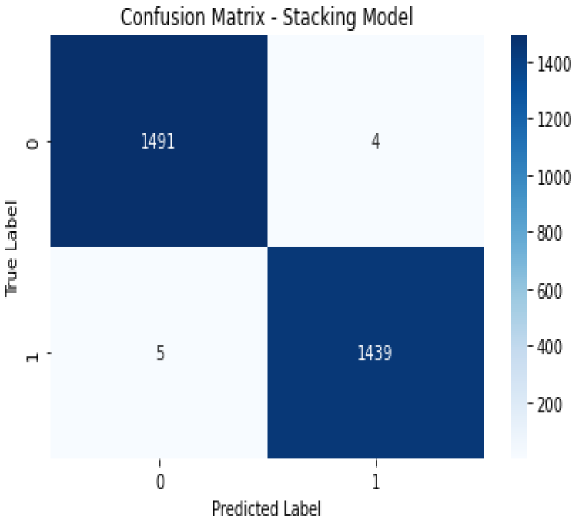


Fig 9. Confusion matrix for the model using stacking after parameter tuning.

4.4 Comparative Results of Voting and Stacking Methods

The performance results of the soft and hard voting ensemble and stacking methods, both with default and tuned parameters, are shown in Table 1.

Table 1. Performance Comparison of the Hybrid Voting Ensemble and Stacking mode.

Model		**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**AUC-ROC (%)**
Default Params	**Soft**	99.15	99.24	99.31	99.27	99.96
	Hard	99.51	99.45	99.38	99.45	-
	Stacking	99.52	99.65	99.65	99.65	99.66
Tuned Params	**Soft**	99.14	99.44	99.24	99.34	99.96
	Hard	99.45	99.45	99.45	99.45	-
	Stacking	99.55	99.65	99.65	99.65	99.66

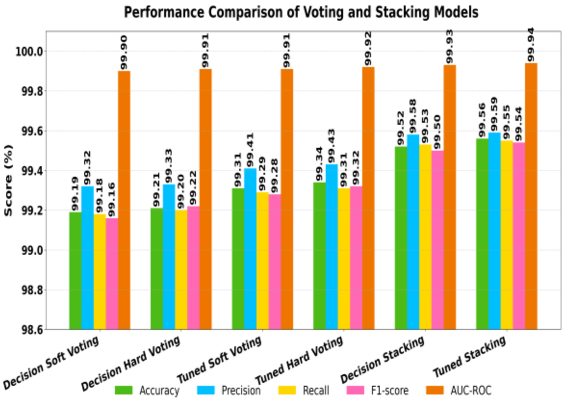


Fig 10. Comparative Analysis of the Results.

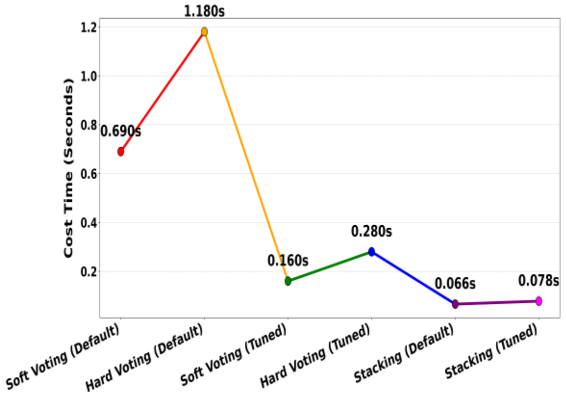


Fig 11. Comparative Analysis of the Cost Time.

As we can see in Figure 11, the execution time of different implemented ensemble models, highlighting their suitability for real-time phishing detection. Among all models, the **stacking approach with tuned parameters** achieved the **fastest response time of just 0.078 seconds**, making it ideal for real-time applications. In contrast, **hard voting with default parameters** had the highest cost time at **1.180 seconds**, indicating slower performance and less practicality in time-sensitive scenarios. Overall, stacking not only offered superior accuracy but also demonstrated efficient processing time, reinforcing its advantage over traditional voting techniques.

CONCLUSIONS

This study presented a hybrid ensemble model for phishing URL detection by integrating of LR, ANN, and RF classifiers. Feature selection was performed using PCA, reducing the input space to 15 optimal features. The model was evaluated using both voting and stacking strategies under default and optimized parameter settings. Results demonstrated that stacking with hyperparameter tuning outperformed other configurations, achieving an accuracy of 99.55% and balanced performance across all evaluation metrics (precision, recall, and F1-score of 99.65%). Furthermore, the optimized stacking model achieved a low prediction time of 0.078 seconds, indicating its effectiveness and efficiency for real-time phishing detection.

4. FUTURE WORK

For future work, employing faster ensemble methods such as XGBoost or LightGBM is recommended to reduce computational costs. In addition, applying optimization techniques like AutoML or evolutionary algorithms may further improve model performance and efficiency.

5. REFERENCES

[1] Phishing URL Detection with ML. https://medium.com/data-science/phishing-domain-detection-with-ml-5be9c99293e5

[2] APWG, "Phishing Activity Trends Report, Quarter4 report 2023," 16 April 2024.

[3] M. Alanezi, "Phishing Detection Methods: A Review," Technium, vol. 3, 2021.

[4] Ali W, Ahmed AA. Hybrid intelligent phishing website prediction using deep neural networks with genetic algorithm-based feature selection and weighting. IET Inf Secur. 2019;13:659–669.

[5] Zhu E, Ye C, Liu D, Liu F, Wang F, Li X. An effective neural network phishing detection model based on optimal feature selection. In: Proceedings of the IEEE International Conference on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications; 2018; Melbourne. p. 781–787.

[6] Mankar NP, Sakunde PE, Zurange S, Date A, Borate V, Mali YK. Comparative evaluation of machine learning models for malicious URL detection. In: Proceedings of the MIT Art, Design and Technology School of Computing International Conference (MITADTSoCiCon); 2024. p. 1–7.

[7] Remeseiro B, Bolon-Canedo V. A review of feature selection methods in medical applications. Comput Biol Med. 2019;112:103375.

[8] Korkmaz M, Sahingoz OK, Diri B. Detection of phishing websites by using machine learning-

based URL analysis. In: Proceedings of the 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT); 2020. p. 1–7.

[9] Laghmati S, Hamida S, Hicham K, Cherradi B, Tmiri A. An improved breast cancer disease prediction system using ML and PCA. Multimed Tools Appl. 2024;83:33785–33821.

[10] Sandunil K, Bennour Z, Ben Mahmud H, Giwelli A. Effects of tuning hyperparameters in random forest regression on reservoir porosity prediction: case study of the Volve oil field, North Sea. In: ARMA US Rock Mechanics/Geomechanics Symposium; 2023. ARMA-2023-0660.

[11] Cosgun AE. Enhancing photovoltaic energy output predictions using ANN and DNN: a hyperparameter optimization approach [Internet]. SSRN; 2024. Available from: https://ssrn.com/abstract=5051166

[12] Yatoo NA, Ali IS. Fine-tuning predictive models: a comprehensive analysis for accurate diabetes risk stratification. Int J Bioinform Res Appl. 2025;21:256–283.

[13] Mridha K, Hasan J, Ghosh A. Phishing URL classification analysis using ANN algorithm. In: Proceedings of the IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON); 2021. p. 1–7.

[14] Das Guptta S, Shahriar KT, Alqahtani H, Alsalman D, Sarker IH. Modeling hybrid feature-based phishing website detection using machine learning techniques. Ann Data Sci. 2024;11:217–242.

[15] Elsheh MM, Swayeb K. Phishing website detection using a hybrid approach based on support vector machine and ant colony optimization. In: Proceedings of the IEEE 3rd International Maghreb Meeting on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA); 2023. p. 402–406.

[16] Karim A, Shahroz M, Mustofa K, Belhaouari SB, Joga SRK. Phishing detection system through hybrid machine learning based on URL. IEEE Access. 2023;11:36805–36822.

[17] Ozcan A, Catal C, Donmez E, Senturk B. A hybrid DNN–LSTM model for detecting phishing URLs. Neural Comput Appl. 2023;35:4957–4973.

[18] Mohanty S, Acharya AA, Gaber T, Panda N, Eldesouky E, Hameed IA. An efficient hybrid feature selection technique toward prediction of suspicious URLs in IoT environment. IEEE Access. 2024;12:50578–50594.

[19] Kocyigit E, Korkmaz M, Sahingoz OK, Diri B. Enhanced feature selection using genetic algorithm for machine-learning-based phishing URL detection. Appl Sci. 2024;14:6081.

[20] Remya S, Pillai MJ, Nair KK, Subbareddy SR, Cho YY. An effective detection approach for phishing URL using ResMLP. IEEE Access. 2024;12:79367–79382.

[21] Javeed MU, Aslam SM, Sadiqa HA, Raza A, Iqbal MM, Akram M. Phishing website URL detection using a hybrid machine learning approach. J Comput Biomed Inform. 2025;9.

[22] Hamidi H, Sayah A. Combining machine learning algorithms to detect phishing URLs: a stacking approach. Int J Eng. 2025;38:1939–1952.

[23] Barik K, Misra S, Mohan R. Web-based phishing URL detection model using deep learning optimization techniques. Int J Data Sci Anal. 2025;1–23.

[24] Prasad YB, Dondeti V. PDSMV3-DCRNN: a novel ensemble deep learning framework for enhancing phishing detection and URL extraction. Comput Secur. 2025;148:104123.

[25] Nayak GS, Muniyal B, Belavagi MC. Enhancing phishing detection: a machine learning approach with feature selection and deep learning models. IEEE Access. 2025.

[26] Phishing detection dataset [Internet]. Kaggle; 2024 Aug 23 [cited 2024 Aug 23]. Available from: https://www.kaggle.com/datasets/eswarchandt/phishing-website-detector

[27] Junsomboon N, Phienthrakul T. Combining over-sampling and under-sampling techniques for imbalance dataset. In: Proceedings of the 9th International Conference on Machine Learning and Computing; 2017. p. 243–247.

[28] Hasan R, Biswas B, Samiun M, Saleh MA, Prabha M, Akter J, et al. Enhancing malware detection with feature selection and scaling techniques using machine learning models. Sci Rep. 2025;15:9122.

[29] Onan A, Korukoğlu S, Bulut H. A multiobjective weighted voting ensemble classifier based on differential evolution

algorithm for text sentiment classification. Expert Syst Appl. 2016;62:1–16.

[30] Ridwan M, Utami E. An optimized hyperparameter tuning for improved hate speech detection with multilayer perceptron. J RESTI (Rekayasa Sistem dan Teknologi Informasi). 2024;8:525–534